



Algorithmic biases: caring about teens' neurorights

José M. Muñoz^{1,2,3} · José Ángel Marinaro⁴

Received: 29 January 2022 / Accepted: 30 May 2022 / Published online: 29 June 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Algorithmic biases are undoubtedly one of the star topics in AI ethics. To get an idea of its importance, we could think of some scandalous recent discoveries in this regard such as a study that showed that the algorithm of the United States health system granted less budget to black people than to white people with the same health problems (Obermeyer et al. 2019). As opposed to this, there is an increasingly influencing trend advocating for the consideration of new human rights, given the ethical challenges that the rapid development of neuroscience and neurotechnology is producing. This view has given rise to proposals of so-called neurorights by groups of philosophers, jurists, and neuroscientists (Ienca and Andorno 2017; Goering et al. 2021). It is among these neurorights that the protection against algorithmic biases occurs. We believe that this is a very interesting and positive initiative. However, the study of algorithmic biases has been focusing mainly on gender and race. Although these two problems seem very serious to us, we want to draw attention to other kind of biases that tend to attract much less attention but affect a particularly vulnerable group: adolescents. The consequences of these biases are especially relevant in criminal law and have particular implications in certain regions of the world such as Latin America.

Traditionally, the foundations of criminal law have been largely based on adult behavior. A clear example is the retributive approach: one is free to choose between complying with the law or not and therefore is worthy of praise

or punishment by virtue of their decision. This approach brings with it the prerequisite of accepting that the individual is mature and free enough to decide, but this is not at all the case with teenagers. There is ample evidence (see, for example, Steinberg 2008) that multiple changes in brain structure and function take place during adolescence and do not end until well into the twenties, which often leads teenagers to make impulsive, dangerous decisions, mostly in the presence of their peers. This phenomenon has been taken into account in various court sentences in the United States—Graham v. Florida, Miller v. Alabama, Roper v. Simmons—thus freeing accused adolescents from severe penalties such as life imprisonment and even execution. By contrast, it is completely ignored in most Latin American countries (Mercurio et al. 2020), where the minimum age of criminal responsibility is set as early as at 12 years—for example, in Mexico and Brazil—or, in the best cases, at 16 years—as in Argentina and Cuba.

But teens' neurobiological immaturity is not only relevant for establishing norms and sentences; it is also closely related to the use of tools for estimating the risk of criminal recidivism. The algorithms used in several of these tools such as COMPAS and LSI-R lead to high risk factors as a result of negatively overestimating records for crimes committed at early ages. This way of estimating risk entirely disregards the dynamics of neurodevelopment and places convicted adolescents in a very vulnerable situation, as it prevents them from having fair access to social reintegration. In addition, with the use of neurotechnologies such as brain imaging for risk estimation—the so-called neuroprediction—this problem of bias could worsen in the coming years. Neurorights proponents argue that “Countermeasures to combat bias should be the norm for algorithms in neurotechnology. Algorithm design should include input from user groups to foundationally address bias” (NeuroRights Foundation 2021). To adequately address this requirement, it is essential to take into account that the integration of data related to activation levels of certain brain areas—amygdala, prefrontal cortex, and cingulate gyrus, among others—in the risk estimation algorithms without keeping in mind that

✉ José Ángel Marinaro
joseangelmarinaro@yahoo.com.ar

¹ Mind-Brain Group, Institute for Culture and Society (ICS), University of Navarra, Pamplona, Spain
² Gordon Center for Medical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States
³ International Center for Neuroscience and Ethics (CINET), Tatiana Foundation, Madrid, Spain
⁴ Department of Law and Political Science, Universidad Nacional de La Matanza, San Justo, Argentina

some of these areas, especially the prefrontal cortex, have not finished developing in adolescents would constitute a new bias that would increase their vulnerability. Nevertheless, as the cognitive traits of adolescents are dynamic—that is, they can be trained and modified—neurotechnology could become a very useful ally for the design and application of specialized intervention programs aimed at achieving their healthy neurobiological development.

While we find that the universal neuroright to protection against algorithmic biases is a convenient general legal principle, we believe that it is also necessary to work on complementary regulatory measures with a view to guaranteeing that this protection is adequately fulfilled in minors. Our specific proposal in this regard is to explore a change in Article 6, Section 2, of the United Nations Convention on the Rights of the Child, which refers to guaranteeing their normal development, by including the right to healthy neurocognitive development and nondiscrimination due to algorithmic and neurotechnological biases. This proposal would harmonize with General Comment No. 24 of the United Nations Committee on the Rights of the Child, of the year 2019, which, in its Sections 22 and 112, claims to consider evidence in the fields of child development and neuroscience that demonstrates their lack of full development and neural maturity. Be that as it may, we believe that this topic should be widely disseminated in national parliaments as well as taken into consideration in judicial decisions, criminal execution, and the determination of the minimum age of criminal responsibility, given its significant impact in these areas. It is critical that the academic community transmits the most relevant research findings to those who make public policy decisions. At the same time, a training policy for justice system operators is urgently needed, considering scholarly findings and debates are not enough if the immense contributions that neuroscience is capable of making are not translated into practice.

E. E. Cummings said, “It takes courage to grow up and become who you really are.” May justice and equal opportunities prevail for those who are yet to unfold their courage.

Acknowledgements José M. Muñoz wants to thank the Tatiana Foundation (<https://fundaciontatianaapgb.org/>) for its valuable support of his research. José Angel Marinero wants to thank the FUNDEJUS (Fundación Estudios para la Justicia; <https://www.fundejus.org>).

Curmudgeon Corner Curmudgeon Corner is a short opinionated column on trends in technology, arts, science and society, commenting on issues of concern to the research community and wider society. Whilst the drive for super-human intelligence promotes potential benefits to wider society, it also raises deep concerns of existential risk, thereby highlighting the need for an ongoing conversation between technology and society. At the core of Curmudgeon concern is the question: What is it to be human in the age of the AI machine? -Editor.

Funding The authors declare that their contribution is framed within Project RTI2018-097838 B-100, funded by MCIN/AEI/10.13039/501100011033/ and FEDER *Una manera de hacer Europa* [<https://blog.uclm.es/proyctodpch/>]. The authors have no relevant financial or non-financial interests to disclose.

References

- Goering S, Klein E, Specker Sullivan L, Wexler A, Agüera y Arcas B, Bi G, Carmena JM, Fins JJ, Friesen P, Gallant J, Huggins JE, Kellmeyer P, Marblestone A, Mitchell C, Parens E, Pham M, Rubel A, Sadato N, Teicher M, Wasserman D, Whittaker M, Wolpaw J, Yuste R (2021) Recommendations for responsible development and application of neurotechnologies. *Neuroethics* 14:365–386. <https://doi.org/10.1007/s12152-021-09468-6>
- Ienca M, Andorno R (2017) Towards new human rights in the age of neuroscience and neurotechnology. *Life Sci Soc Policy* 13:5. <https://doi.org/10.1186/s40504-017-0050-1>
- Mercurio E, García-López E, Morales-Quintero LA, Llamas NE, Marinero JA, Muñoz JM (2020) Adolescent brain development and progressive legal responsibility in the Latin American Context. *Front Psychol* 11:627. <https://doi.org/10.3389/fpsyg.2020.00627>
- NeuroRights Foundation (2021) Mission. <https://neurorightsfoundation.org/mission>
- Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366(6464):447–453. <https://doi.org/10.1126/science.aax2342>
- Steinberg L (2008) A social neuroscience perspective on adolescent risk-taking. *Dev Rev* 28(1):78–106. <https://doi.org/10.1016/j.dr.2007.08.002>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.